Image Enhancement and Automated Target Recognition Techniques for Underwater Electro-Optic Imagery

Thomas Giddings (PI), Cetin Savkli and Joseph Shirron
Metron, Inc.
11911 Freedom Dr., Suite 800
Reston, VA 20190
phone: (703) 437-2428 fax: (703) 787-3518 email: giddings@metsci.com

Contract Number N00014-07-C-0351
http:www.metsci.com

LONG TERM GOALS

The long-term goal of this project is to provide a flexible, accurate and extensible automated target recognition (ATR) system for use with a variety of imaging and non-imaging sensors. Such an ATR system, once it achieves a high level of performance, can relieve human operators from the tedious business of pouring over vast quantities of mostly mundane data, calling the operator in only when the computer assessment involves an unacceptable level of ambiguity. The ATR system will provide most leading edge algorithms for detection, segmentation, and classification while incorporating many novel algorithms that we are developing at Metron. To address one of the most critical challenges in ATR technology, the system will also provide powerful feature extraction routines designed for specific applications of current interest.

OBJECTIVES

The main objective of this project is to develop a complete, flexible, and extensible modular automated target recognition (MATR) system for computer aided detection and classification (CAD/CAC) of target objects from within cluttered and possibly noisy image data. The MATR system framework is designed to be applicable to a wide range of situations, each with its own challenges, and so is organized in such a way that the constituent algorithms are interchangeable and can be selected based on their individual suitability to the particular task within the specific application. The ATR system designer can select combinations of algorithms, many of which are being developed at Metron, to produce a variety of systems, each tailored to specific needs. While the development of the system is still ongoing, results for mine countermeasures (MCM) applications using electro-optical (EO) image data have been encouraging. A brief description of the system framework, some of the novel algorithms, and preliminary test results are provided in this interim report.

APPROACH

The MATR system is composed of several modules, as depicted in Figure 1, reflecting the sequence of steps in the ATR process. The detection step is concerned with finding portions of an image that contain possible objects of interest, or targets, that merit further attention. During the localization and segmentation phase the position and approximate size and shape of the object is estimated and a portion of the image, or "snippet," containing the object is extracted. At this stage, image processing may be performed on the snippet to reorient the target, mitigate noise, accentuate edge detail, etc.

Low-level pixel information (e.g., pixel intensities, hues, etc.) is then arranged into an attribute vector. At the feature extraction step, the most discriminating high-level characteristics of the attribute vectors are distilled from the attribute vector, and this summary information is formed into the feature vector. The feature vector is submitted to the classification process, where the object in the snippet is classified or identified. Performance of the system is typically measured in terms of probabilities of correct classification/identification and false alarm rates.

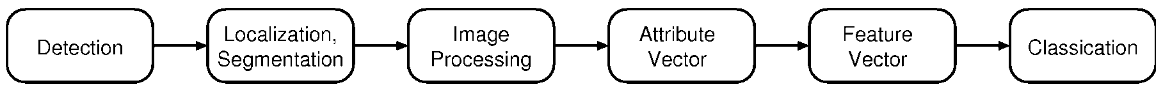

Figure 1: Schematic of modular automated target recognition system.

While the MATR system is designed to address many computer assisted target detection and classification problems using a different types sensor data, it is presently being applied to bottom mine hunting using electro-optical imaging systems such as those incorporated within the AN/AQS-24 and AN/AQS-20 systems. Electro-optical system imagery used for MATR system development and testing is taken from a data set acquired during EOID systems tests, sponsored by the Office of Naval Research (ONR), which took place in the Panama City Beach (PCB) area in August 2001 [1]. Selected image snippets of various targets and clutter objects are shown in Figure 2 and Figure 3, respectively. These objects display a variety of shapes and patterns, but there are similarities between some of them. For example, the Manta mine shape and the tire are similar in shape, and the DST-36 mine shape resembles the shell casing. The snippets in these figures are some of the clearest examples of each object in the data set, while many other snippets are occluded by fish schools, degraded by noise, and/or marked by very low contrast. For example, at the top of the images in Figure 4 a school of fish is obstructing the view of a culvert lying on the bottom.

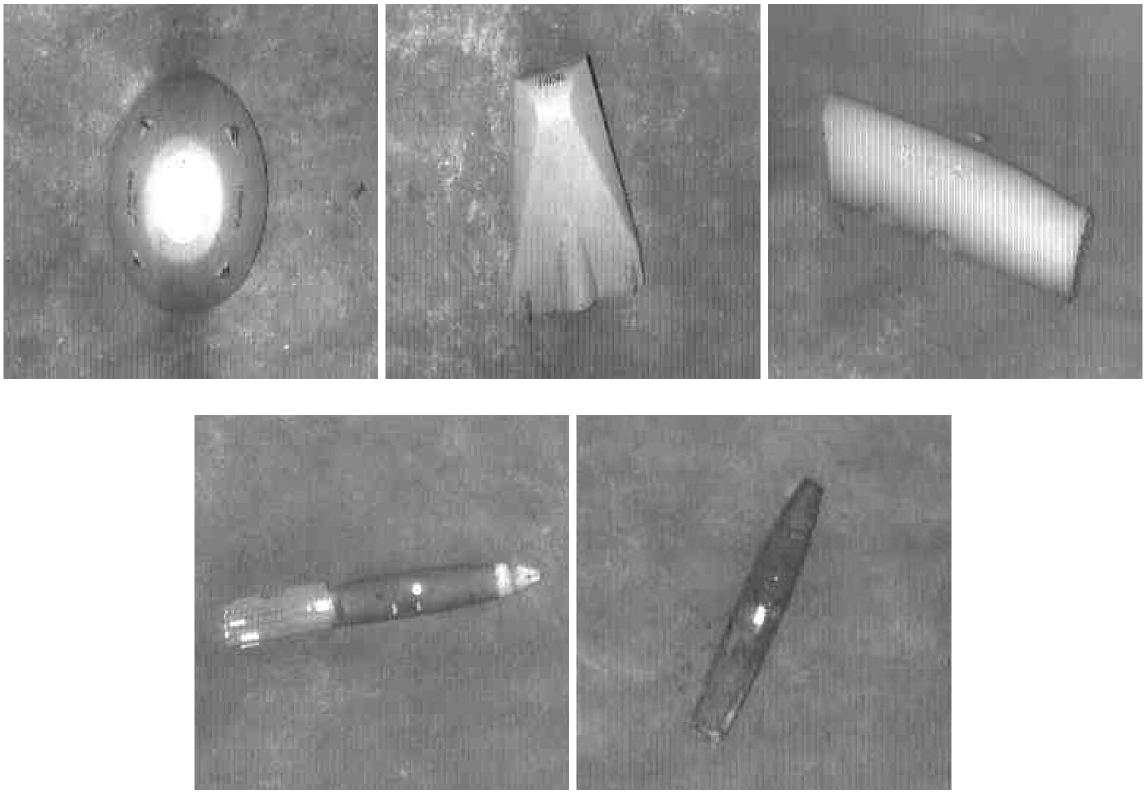

Figure 2: Example snippets of various targets in field. Clockwise from the top left are a Manta mine shape, a Rockan, an MR-80, a shell casing, and a DST-36.

Other data sets that are being used in the development of the MATR system include synthetic aperture sonar imagery from the Small Synthetic Aperture Minehunter (SSAM), collected during the AUV Fest in 2005, and the Yale Face databases [2]. The synthetic aperture sonar data is mainly being used to support the development of object detection and segmentation algorithms. The Yale Face set of images is being used to develop and test the feature extraction and classification algorithms. The use of data from a variety of different sources ensures that the methodologies being developed are robust, and are not relying on information particular to a specific sensor or situation. Also, it is important to note that the Yale Face images are all oriented in the same way while the target and clutter data from the EOID imagery is currently submitted to the classifier without reorienting the objects to line up in any particular way. The lack of consistent orientation in the EOID image snippets complicates matters for the classifier. Perhaps surprisingly, with sufficient training data from which to learn the various orientations the classifier performance is still very good with the EOID dataset without reorientation, but will no doubt improve with the development of a reorientation algorithm.

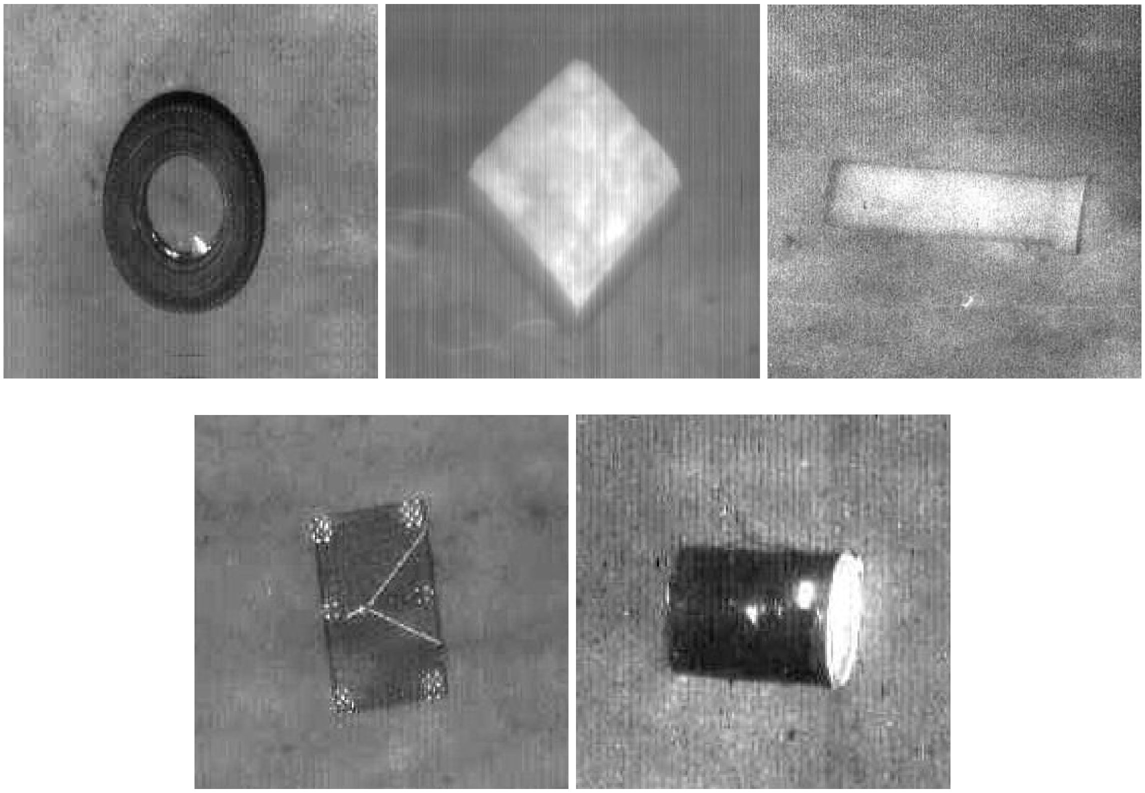

Figure 3: Example snippets of various clutter objects in field. Clockwise from the top left are a tire, a concrete clump, a culvert, an oil drum, and a crab trap.

WORK COMPLETED

The following is a brief description of current algorithms in the various ATR modules. Comparisons with other state-of-the-art approaches are demonstrated. Classification methodology and results are also described.

Object Detection and Segmentation

Detection involves locating objects within an image. Segmentation requires further knowledge of the approximate sizes and shapes of the objects. In the present case, we assume no prior knowledge of location, object size or shape.. For grayscale imagery we propose a simple detection/segmentation technique based on the histograms of image subframes. Within the subframe the image is reduced to a binary mask based on whether the pixel intensities are greater or less than the mean intensity within the frame. The resulting binary mask is then compared to a circle inscribed within the frame and assigned a rating according to a scaled inner product. A rating surpassing a threshold indicates a "fit," or detection. In Figure 4 we show a typical comparison between our "histogram" technique and a relatively new method described in Ref. [3-5], which uses image entropy maximization to determine the salient regions of an image,. The results clearly demonstrate the superiority of the "histogram" technique.

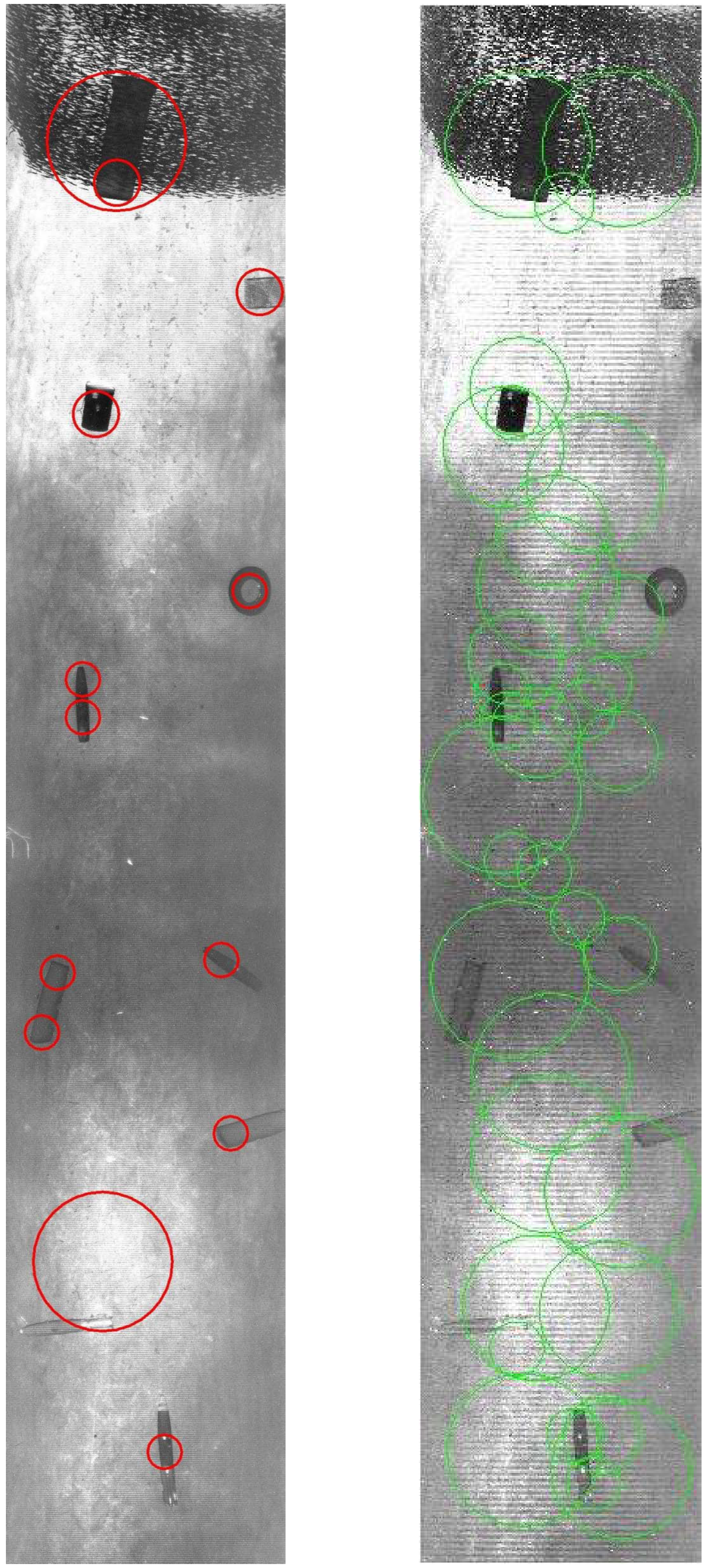

Figure 4: Comparison of the "histogram" detection method (left) with the entropy maximization approach (right) using EOID data.

Image Processing

EOID systems are often power limited, so that low signal-to-noise ratio (SNR) typically plagues EOID imagery. This type of image degradation can be ameliorated by using median or mode noise reduction filters, which are described in Ref. [6], for example. These filters work well at reducing noise while preserving texture and edge detail, which tend to get blurred away by Gaussian noise reduction filters. Figure 5 shows a checkerboard pattern with varying length scales with and without additive noise. Figure 6 shows the results after several passes of median, mode, and Gaussian de-noising filters. It is clear that the texture at smaller scales is blurred away by the Gaussian filter while it is still discernable (though degraded) with the median and mode filters. Image enhancement techniques such as these de-noising algorithms are an important preprocessing step that we apply prior to feature extraction and classification, where it is important not to corrupt identifying image features such as texture.

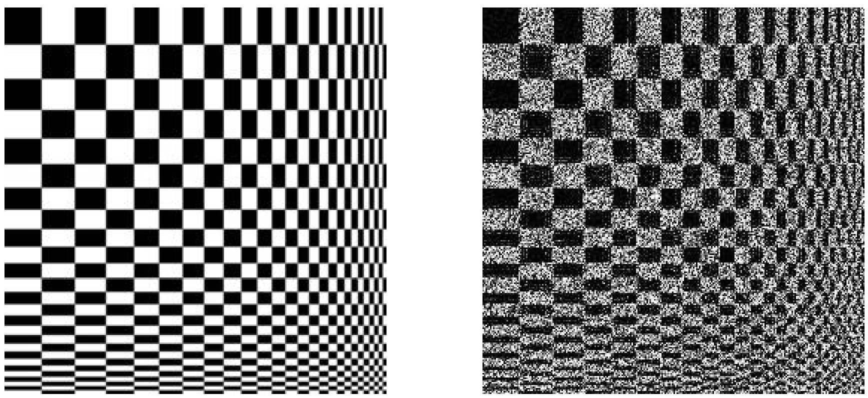

Figure 5: Checkerboard image (left) and the same image with additive Guassian noise (right)

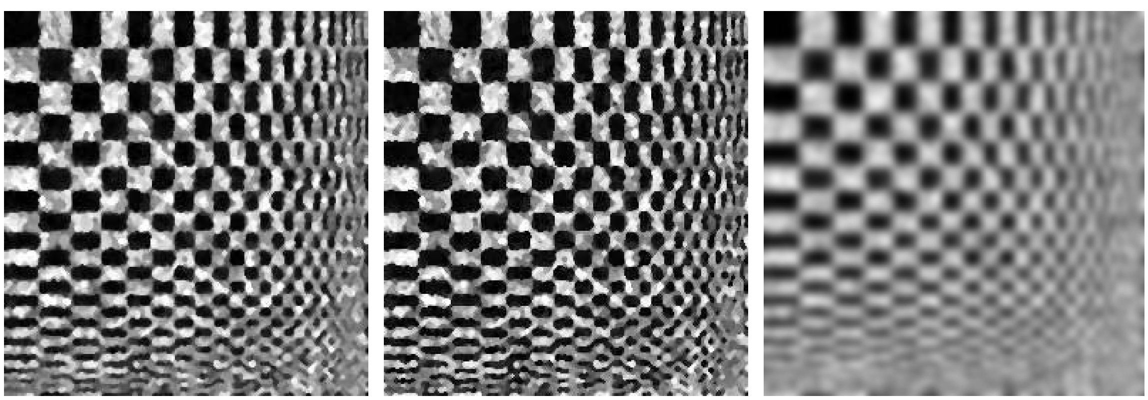

Figure 6: Noisy image after several iterations of a median filter (left), mode filter (middle) or Gaussian filter (right) and contrast stretching.

Feature Extraction

Image snippets generated by the segmentation process can be composed of many thousands, tens of thousands, or even millions of pixels. For classification purposes we first extract image features using the Principle Components Analysis (PCA) method developed in Ref. [7]. The fundamental objective of PCA is to reduce the dimensionality of the problem by expanding the image data in terms of the eigenmodes, and then retaining only the modes that are a rich source of identifying information [8]. The eigenmodes constitute a new basis for the image description, and most of the class-to-class variability in the images, and therefore most of the image relevant information, is contained in the directions of the eigenmodes associated with the largest eigenvalues. By using PCA we can reduce the dimensionality of the image description by several orders of magnitude, thereby speeding up the classification process, while maintaining the same level of classifier performance.

Classification Method and Results

A variety of classification methods can be employed. We are currently employing a neural network (NN) to perform the classification of the image vectors. We have tested our approach using the Yale Face database, which contains 15 subjects with 11 images each. We used 10 for training and 1 image for testing for each subject, in various combinations, with perfect results (100% correct identification). By way of comparison, a Euclidean distance based classification scheme achieved a lower performance of 90% correct identification.

A similar training and testing procedure was used on a second, larger database of underwater electro-optical bottom target imagery, examples of which are shown in Figures 2 and 3. This database demonstrated several complicating factors such as variations in orientation and scale within the image snippets. Despite these disadvantages the PCA+NN approach correctly identified approximately 90% of the targets in these test images. Again, the Euclidean distance based classification scheme posted a poorer performance with about 80% of the targets correctly identified.

RESULTS

Main results include the development of several novel object detection and segmentation techniques that outperform other documented methods. These novel techniques are currently being extended to accurately determine the shape and boundaries of contiguous objects. Further, a very powerful and general technique combining feature extraction with neural network classifiers has been developed for 2D grayscale images and which provides excellent results across a range of applications (e.g., face recognition, mine identification, etc.) . The approach taken provides excellent classification results across a variety of applications.

IMPACT/APPLICATIONS

The two most important remaining problems in automated target recognition are segmentation and feature extraction. With the development of fast, accurate and robust segmentation algorithms, capable of delineating the boundaries of complex patterned objects against various backgrounds, real-world application of ATR systems to MCM operations becomes much more feasible. Classifier technology being very mature, the performance of such systems then hinges mainly on the quality of the features that are extracted from the data segment and submitted to the classifier. Ideally, a high-performance ATR system would impact MCM operations by replacing the human operator altogether. More realistically, an effective ATR system would elevate the human operator to a supervisory role where the computer takes on the more mundane tasks that otherwise would occupy the vast majority of the operator's time. This would reduce manpower requirements and eliminate the errors due to human operator fatigue.

RELATED PROJECTS

None.

REFERENCES

[1] J. S. Taylor Jr. and M. C. Hulgan, "Electro-Optic Identification Research Program," in *MTS/IEEE Oceans '02 Conf. Proc.*, vol. 2, pp. 994-1002, 2002.

[2] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose," *IEEE Trans. Pattern Anal. Mach. Intelligence*, vol. 23, no. 6, pp. 643-660, 2001.

[3] T. Kadir and M. Brady, "Scale, saliency and image description," *Intl. J. of Computer Vision*, vol.. 45, no. 2, pp. 83-105, 2001.

[4] T. Kadir and M. Brady, "Scale saliency: A novel approach to salient feature and scale selection," in *Intl. Conf. on Visual Information Engineering (VIE 2003)*, Guildford, UK, pp. 25-28, 2003.

[5] T. Kadir, A. Zisserman, and M. Brady, "An affine invariant salient region detector," in *Proceedings of the 8th European Conference on Computer Vision*, Prague, Czech Republic, pp. 228-241, 204.

[6] M. Nixon and A. Aguado, *Feature Extraction & Image Processing*, Newnes, 2002.

[7] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, vol. 3, no. 1, pp.71-86, 1991.

[8] I. T. Jolliffe, *Principle Component Analysis*, New York, NY, Springer-Verlag, 2nd edition, 2004.